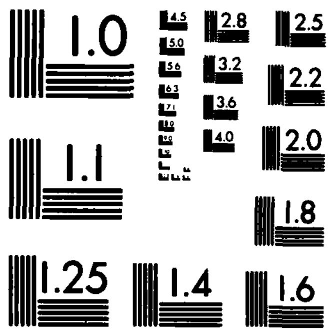
AD-A158 869   A GRAPHICAL SIMILARITY MEASURE FORTTIME SERIES MODELS   1/1
(U) MARYLAND UNIV COLLEGE PARK DEPT OF MATHEMATICS
B KEDEM APR 85 TR-85-10 AFOSR-TR-85-0680 AFOSR-82-0187

UNCLASSIFIED   F/G 12/1   NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

(2)

AD-A158 869

REPORT DOCUMENTATION PAGE

1b. RESTRICTIVE MARKINGS	
3. DISTRIBUTION/AVAILABILITY OF REPORT	Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	University of Maryland Mathematics Department
5. MONITORING ORGANIZATION REPORT NUMBER(S)	AFOSR-TR- 85-0686
6a. NAME OF PERFORMING ORGANIZATION	College Park, MD. 20742
6b. OFFICE SYMBOL (If applicable)	
7a. NAME OF MONITORING ORGANIZATION	Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code)	
7b. ADDRESS (City, State and ZIP Code)	Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	AFOSR
8b. OFFICE SYMBOL (If applicable)	NM
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	AFOSR 82-0187
8c. ADDRESS (City, State and ZIP Code)	Bolling AFB DC 20332-6448

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE (Include Security Classification)
A GRAPHICAL SIMILARITY MEASURE FOR TIME SERIES MODELS

12. PERSONAL AUTHOR(S)
Benjamin Kedem

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Tech. Report	FROM ____ TO ____	April 1985	21

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Oscillation, Spectral, Discrimination, Classification, Goodness of fit, HOC

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The rate of increase of higher order crossings (HOC) is a powerful discriminatory device. A new approximation for the variance of HOC is obtained and used in the construction of probability limits for HOC. Uses of these limits for model selection are illustrated via real and simulated data.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Major Brian Woodruff	(202) 767-	NM

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- 85-0680

A Graphical Similarity Measure For Time Series Models

Benjamin Kedem

University of Maryland
College Park, MD.

Accession For		
NTIS GRA&I		X
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A-1		

TR85-10
MD85-11-BK

85 9 05 114

A GRAPHICAL SIMILARITY MEASURE FOR TIME SERIES MODELS

Benjamin Kedem
University of Maryland, College Park

Abstract

The oscillation in time series can be represented by counts of axis-crossings in the series and its differences. These counts are called higher order crossings and display a monotone property whose rate of increase discriminates between processes. Only very few of these counts are needed for effective discrimination as shown by plots of higher order crossings obtained from real and simulated data.

Key Words: Oscillation, Spectral, Probability limits, Axis-crossings, Stationary.

The support of AFOSR Grant 82-0187 is gratefully acknowledged.

1. Introduction

The purpose of this note is to introduce a graphical device useful as a measure of similarity or as a goodness of fit criterion for hypothesized time series models. It is based on the actual oscillation observed in time series as depicted by axis-crossings and higher order crossings. Higher order crossings (HOC) are axis-crossings of differenced time series and are closely linked to the spectral content of the series. In fact under the Gaussian assumption, to which we shall adhere, HOC determine the finite dimensional distributions up to a scale parameter given that the mean is zero. The main advantage of HOC is that they are easily obtained from an observed series and that only very few of them are needed, as the discriminatory power in HOC usually diminishes with their order.

Higher order crossings in time series discrimination were discussed in Kedem and Slud (1981), (1982), where a certain goodness of fit criterion is suggested. Here however the emphasis is on a graphical device rather than a single test statistic. This graphical method may be shown useful in answering the question "Does a given time series oscillate as a certain hypothesized model?" Some examples with real and simulated data demonstrate the use and potential of this method.

2. Definition and Properties of HOC

Let $\{Z_t\}$, $t = 0, \pm 1, \ldots,$ be a zero mean stationary Gaussian process with correlation function ρ_j and spectral distribution function F. The higher order crossings of order k, $D_{k,N}$, is the number of axis-crossings by

$$\nabla^{k-1} Z_1, \ldots, \nabla^{k-1} Z_N$$

where ∇ is the difference operator

$$\nabla Z_t = Z_t - Z_{t-1}$$

and $\nabla^k Z_t = \nabla(\nabla^{k-1} Z_t)$. Since ∇^0 is the identity operator, $D_{1,N}$ is the actual number of axis-crossings by $Z_1, \ldots, Z_N$.

A more precise definition is constructed in terms of

$$X_t^{(k)} \equiv \begin{matrix} 1, & \nabla^{k-1} Z_t \geq 0 \\ 0, & \text{otherwise} \end{matrix}, \quad k = 1, 2, \ldots$$

and

$$d_t^{(k)} \equiv \begin{matrix} 1, & X_t^{(k)} \neq X_{t-1}^{(k)} \\ 0, & \text{otherwise.} \end{matrix}$$

We define formally the higher order crossings of order k by

$$D_{k,n} \equiv d_2^{(k)} + \ldots + d_N^{(k)}.$$

In practice we only deal with finite series and some care must be taken in regard to the differencing of finitely long records. Each time a finite series is differenced an observation is lost so that if k higher crossings are desired the time index $t = 1$ is given to the k'th or a later observation and we shall follow this rule.

Consider the following example. Suppose the series Z_t is given by

$$-1 \quad .4 \quad -3 \quad .6 \quad -.3 \quad -.8 \quad 1.2 \quad -1.3 \quad -.4 \quad -1.1.$$

In order to derive the first three HOC $D_{1,6}$, $D_{2,6}$, $D_{3,6}$ we have

t					1	2	3	4	5	6
Z	-1	.4	- .3	.6	- .3	-.8	1.2	-1.3	-.4	-1.1
$X^{(1)}$	0	1	0	1	0	0	1	0	0	0
∇Z		1.4	- .7	.9	- .9	-.5	2.0	-2.5	.9	- .7
$X^{(2)}$		1	0	1	0	0	1	0	1	0
$\nabla^2 Z$			-2.1	1.6	-1.8	.4	2.5	-4.5	3.4	-1.6
$X^{(3)}$			0	1	0	1	1	0	1	0

and $D_{1,6} = 2$, $D_{2,6} = 4$, $D_{3,6} = 4$ and we say that D_1, D_2, D_3 were derived from a series of length 6. It is seen that counting symbol changes in the clipped binary series is equivalent to counting axis-crossings in the corresponding series. Thus $D_{1,6} = 2$ since there are 2 symbol changes in $X_1^{(1)},\ldots,X_6^{(1)}$ or equivalently 2 axis-crossings in $Z_1,\ldots,Z_6$. It is interesting to observe that while $D_{1,6}$ is the number of crossings, $D_{2,6}$ is the number of peaks and troughs in $Z_1,\ldots,Z_6$.

The fact that the correlation structure in $\{Z_t\}$ is completely determined by the sequence

$$\{ED_{j,N}\}_{j=1}^{\infty}$$

is due to the basic relation

$$\rho_1 = \cos\left(\frac{\pi ED_{1,N}}{N-1}\right) \tag{1}$$

whose extension is given in (2) by considering the first correlation in $\nabla^k Z_t$,

$$\cos\left(\frac{\pi ED_{k+1,N}}{N-1}\right) = \frac{-\binom{2k}{k-1} + \rho_1\left\{\binom{2k}{k} + \binom{2k}{k-2}\right\} - \dots + (-1)^k \rho_{k+1}}{\binom{2k}{k} - 2\rho_1\binom{2k}{k-1} + \dots + (-1)^k 2\rho_k}. \tag{2}$$

From (1) and (2) we can determine $\rho_1, \rho_2, \rho_3, \dots$ recursively. (See Kedem and Slud (1981).) (2) can also be written as

$$\cos\left(\frac{\pi ED_{k+1,N}}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega)(\sin \tfrac{1}{2}\omega)^{2k} dF(\omega)}{\int_{-\pi}^{\pi} (\sin \tfrac{1}{2}\omega)^{2k} dF(\omega)} \tag{3}$$

which is referred to as the higher order crossings spectral representation and relates the sequence $\{ED_{j,N}\}$ to F. This spectral representation has been recently studied in Kedem (1984) in some detail, where it was shown that

$$ED_{1,N} \le ED_{2,N} \le \dots \le (N-1). \tag{4}$$

It follows that $\{ED_{j,N}\}_{j=1}^{\infty}$ is a <u>monotone increasing and bounded</u> sequence and therefore converges. It can be shown that

$$\frac{\pi ED_{j,N}}{N-1} \to \omega^*, \quad j \to \infty, \tag{5}$$

where ω^* is the highest point in the support of F, that is, the highest frequency in the spectrum. Therefore

$$ED_{1,N} \le ED_{2,N} \le \dots \le \frac{\omega^*}{\pi}(N-1). \tag{6}$$

In practice most series contain noise and $\omega^* = \pi$. When $\omega^* = \pi$ the inequalities in (4), (6) are strict. It was shown in Kedem and Slud (1982) that when π is in fact included in the support of F then the actual sequence $\{D_{j,N}\}$ is monotone increasing provided N is sufficiently large.

Now it is this monotone property of HOC which we would like to utilize. More precisely, we wish to demonstrate that the initial rate of increase in the $D_{j,N}$ as $j \to \infty$ is rather fast and carries discriminatory information as different processes display different rates of increase in the $D_{j,N}$ for small j and fixed N. As j increases this discriminatory potency decreases rapidly and similar rates are displayed by many different series. In this sense only very few $D_{j,N}$ are needed in achieving effective discrimination.

It should be noted that from (2), (3), the information contained in $\{ED_{j,N}\}_{j=1}^{\infty}$ is equivalent to that in $\{\rho_j\}$ and consequently that in $\dot{F}$, the normalized spectral distribution function. But the information manifested by the $ED_{j,N}$ presents a different angle of view and another way for the interpretation of stationary time series. HOC provide another dimension in time series analysis, which addresses the oscillatory content in time series directly by counts of visual features.

3. An Approximation to the Variance of HOC

As we shall restrict our attention to the rate of increase in HOC it is necessary to determine beforehand the variance of $D_{k,N}$, $k = 1,2,\ldots$. In general this variance is a function of 4'th order orthant probabilities which are not easily accessible (see Reed (1983)). It has been observed however that in many cases the dependence in the $\{d_t^{(k)}\}$ for low k is rather weak as expressed by very low correlations, unless the original series $\{Z_t\}$ displays extremely large absolute correlations. Thus as a first approximation we suggest treating $\{d_t^{(k)}\}$ as a binary Markov chain. This simplification leads immediately to a computable expression for $\mathrm{Var}(D_{k,N})$ which, as we shall see, provides a rather close approximation to the true variance.

Define

$$\lambda_j^{(k)} \equiv P(X_t^{(k)} = 1 \mid X_{t-j}^{(k)} = 1)$$

$$\rho_j^{(k)} \equiv \mathrm{Corr}(\nabla^k Z_t \, , \nabla^k Z_{t-j}).$$

Then from (1) or (2) we have

$$\lambda_j^{(k)} = \frac{1}{2} + \frac{1}{\pi} \sin^{-1}\rho_j^{(k-1)}. \tag{7}$$

Observe that because of (1), $\rho_1^{(k)}$ is precisely given by (2). The two parameters which specify $\{d_t^{(k)}\}$ are thus given by

$$p^{(k)} \equiv P(d_t^{(k)} = 1) = 1 - \lambda_1^{(k)} \tag{8}$$

$$\nu^{(k)} = P(d_t^{(k)} = 1 \mid d_{t-1}^{(k)} = 1) = \frac{1 - 2\lambda_1^{(k)} + \lambda_2^{(k)}}{2(1 - \lambda_1^{(k)})}. \tag{9}$$

Therefore using standard arguments from the theory of Markov chains (see Karlin and Taylor (1975), Ch. 2) we have the approximation owing to the fact that $D_{k,N} = d_2^{(k)} + \ldots + d_N^{(k)}$,

$$\mathrm{Var}(D_{k,N}) = (N-1)p^{(k)}q^{(k)} + \frac{2p^{(k)}q^{(k)}(\nu^{(k)} - p^{(k)})}{(1-\nu^{(k)})}\left\{(N-1) - V_{k,N}\right\} \qquad (10)$$

where

$$V_{k,N} = q^{(k)}\left[1 - \left(\frac{\nu^{(k)} - p^{(k)}}{q^{(k)}}\right)^{N-1}\right] \Big/ (1-\nu^{(k)})$$

and $q^{(k)} = 1 - p^{(k)}$. Usually $V_{k,N}$ is negligible compared with N and may be omitted. Thus, a useful approximation to $\mathrm{Var}(D_{k,N})$ is

$$\mathrm{Var}(D_{k,N}) \simeq (N-1)p^{(k)}q^{(k)}\left[\frac{1-2p^{(k)}+\nu^{(k)}}{1-\nu^{(k)}}\right]. \qquad (11)$$

From a given sequence $\rho_1, \rho_2, \ldots, \rho_K$ we compute $\mathrm{Var}(D_{k,N})$, $k = 1, \ldots, K-1$ as follows.

(i) Obtain $\rho_1^{(k)}$ from (2), $k = 0, 1, \ldots, K-1$.

$$\rho_1^{(k)} = \frac{-\binom{2k}{k-1} + \rho_1\left\{\binom{2k}{k} + \binom{2k}{k-2}\right\} - \rho_2\left\{\binom{2k}{k-1} + \binom{2k}{k-3}\right\} + \ldots + (-1)^k \rho_{k+1}}{\binom{2k}{k} - 2\rho_1\binom{2k}{k-1} + \ldots + (-1)^k 2\rho_k}.$$

Then $\rho_2^{(k)}$ is given by

$$\rho_2^{(k)} = -1 + 2\rho_1^{(k)} - 2\rho_1^{(k+1)}(1-\rho_1^{(k)}).$$

(ii) Obtain $\lambda_1^{(k)}$, $\lambda_2^{(k)}$ from (7), $k = 1, \ldots, K-1$.

$$\lambda_1^{(k)} = \frac{1}{2} + \frac{1}{\pi}\sin^{-1}\rho_1^{(k-1)}, \qquad \lambda_2^{(k)} = \frac{1}{2} + \frac{1}{\pi}\sin^{-1}\rho_2^{(k-1)}.$$

(iii) Obtain $P^{(k)}$, $\nu^{(k)}$, from (8), (9), $k = 1,\ldots,K-1$

$$P^{(k)} = 1 - \lambda_1^{(k)}, \qquad \nu^{(k)} = \frac{1 - 2\lambda_1^{(k)} + \lambda_2^{(k)}}{2(1 - \lambda_1^{(k)})} .$$

(iv) Substitute in (10) to obtain $\mathrm{Var}(D_{k,N})$, $k = 1,\ldots,N-1$.

The above algorithm results are compared in what follows with actual estimates in the estimation of $\mathrm{Var}(D_{j,1000})$, $j = 1,\ldots,6$ for several series. To apply the algorithm one needs to supply as input only $\rho_1,\ldots,\rho_7$ and N. (A computer program which takes as input $\rho_1,\ldots,\rho_7,N$ and gives as output $E(D_{j,N})$ and $\mathrm{Var}(D_{j,N})$ is available upon request.) When an hypothesized model is entertained these parameters are usually available as is the case for example when the hypothesized process is white noise or any specific autoregressive moving average. The experimental results were obtained from 100 independent realizations of size $N = 1000$ each and are summarized in Tables 1, 2, 3. Note that $ED_{j,1000}$ is available exactly and does not depend on the Markov assumption. The tables also compare $ED_{j,1000}$ with their estimates as a check on the simulation results. It is seen that $ED_{j,1000}$ and $\hat{E}D_{j,1000}$ are very close throughout. In particular, $\{\mathrm{Var}\ D_{j,1000}\}^{1/2}$ obtained from (10) agrees well with actual estimates obtained from these 100 realizations. We note that throughout our simulation $V_{k,N}$ in (10) ranged in value from 0.663 to 1.247 but most often its value was around 0.9, which is very small compared with $N = 1000$. In specifying a model we use the notation as in Box and Jenkins (1970). For example, AR(1) with parameter ϕ means $Z_t = \phi Z_{t-1} + u_t$, but MA(1) with parameter θ means $Z_t = u_t - \theta u_{t-1}$, where u_t are $N(0,\sigma_u^2)$ independently distributed. It should be noted that all our results concerning HOC are scale free so that the actual magnitude of σ_u^2 is not important.

FIRST ORDER MODELS

Series	j	$ED_{j,1000}$	$\hat{ED}_{j,1000}$	$\{Var\ D_{j,1000}\}^{\frac{1}{2}}$ from (10)	$\{\widehat{Var}\ D_{j,1000}\}^{\frac{1}{2}}$ from 100 realizations
White Noise	1	500	497	15.81	15.96
	2	666	666	13.15	13.63
	3	732	732	12.16	12.53
	4	769	770	11.57	11.49
	5	794	795	11.18	11.05
	6	813	814	10.82	10.00
AR(1) $\phi = 0.5$	1	333	333	15.74	15.58
	2	580	582	14.00	15.03
	3	685	686	12.72	13.11
	4	740	740	12.00	12.08
	5	774	774	11.49	11.23
	6	797	796	11.09	10.00
MA(1) $\theta = -0.5$	1	369	371	14.17	14.59
	2	553	551	12.44	12.27
	3	640	639	11.91	12.48
	4	694	693	11.66	12.22
	5	732	731	11.44	11.40
	6	759	759	11.22	11.02

Table 1. Comparison of $\{Var\ D_{j,1000}\}^{\frac{1}{2}}$ from (10) and the estimate obtained from 100 independent realizations each of size 1000 given in the last column. $ED_{j,1000}$ are rounded to the nearest integer.

SECOND ORDER MODELS

Series	j	$ED_{j,1000}$	$\hat{ED}_{j,1000}$	$\{\text{Var } D_{j,1000}\}^{1/2}$ from (10)	$\{\widehat{\text{Var } D_{j,1000}}\}^{1/2}$ from 100 realizations
AR(2)	1	333	333	12.00	10.83
$\phi_1 = 0.75$	2	460	460	11.87	11.07
	3	562	562	12.40	12.58
$\phi_2 = -0.5$	4	646	649	12.65	13.58
	5	708	709	12.25	12.57
	6	751	751	11.79	11.91
MA(2)	1	342	343	18.33	17.76
$\theta_1 = -0.5$	2	648	649	17.66	18.06
	3	782	784	13.82	13.71
$\theta_2 = -0.8$	4	835	836	11.40	10.49
	5	857	858	10.24	9.56
	6	868	869	9.74	8.71
AR(2)	1	424	425	9.64	9.67
$\phi_1 = 0.4$	2	484	485	9.38	9.13
	3	536	537	10.29	10.81
$\phi_2 = -0.7$	4	594	594	11.27	12.72
	5	651	652	11.87	12.02
	6	702	701	12.04	11.34

Table 2. Comparison of $\{\text{Var } D_{j,1000}\}^{1/2}$ from (10) with the estimate obtained from 100 independent realizations of size 1000 given in the last column. $ED_{j,1000}$ are rounded to the nearest integer.

MIXED MODELS

Series	j	$ED_{j,1000}$	$\hat{ED}_{j,1000}$	$\{\text{Var } D_{j,1000}\}^{1/2}$ from (10)	$\{\widehat{\text{Var}}\ D_{j,1000}\}^{1/2}$ from 100 realizations
ARMA(1,1)	1	552	552	14.62	14.74
$\phi = 0.5$	2	679	679	12.96	12.87
	3	737	737	12.09	12.05
$\theta = 0.7$	4	773	772	11.27	11.52
	5	797	797	10.70	11.12
	6	814	814	10.15	10.80
ARMA(1,1)	1	235	234	13.17	12.67
$\phi = 0.5$	2	432	433	11.59	12.07
	3	544	544	11.42	11.25
$\theta = -0.7$	4	615	615	11.04	11.07
	5	663	665	11.26	11.05
	6	698	699	10.65	11.03
ARMA(2,2)	1	884	883	10.04	10.51
$\phi_1 = -1.4$	2	897	897	9.20	9.53
$\phi_2 = -0.5$	3	903	903	8.84	9.01
$\theta_1 = 0.2$	4	908	908	8.60	8.50
$\theta_2 = 0.1$	5	911	911	8.43	8.47
	6	914	914	8.29	8.38

Table 3. Comparison of $\{\text{Var } D_{j,1000}\}^{1/2}$ from (10) with the estimate obtained from 100 independent realizations of size 1000 given in the last column. $ED_{j,1000}$ are rounded to the nearest integer.

4. Probability Limits for HOC

The asymptotic distribution of $D_{k,N}$, as $N \to \infty$, can be found under standard moment conditions without recourse to the Markov assumption which we only needed in getting an approximation to the variance of HOC. For any finite moving average series $\{Z_t\}$, $\{d_t^k\}$ is m-dependent for some finite m which implies (Diananda (1953)) that the asymptotic distribution of $D_{k,N}$ as $N \to \infty$ is normal. More generally, since $\{Z_t\}$ is Gaussian the condition $\sum |\rho_k| < \infty$ implies the asymptotic normality of $D_{k,N}$ for any finite k. The proof of this fact uses a method suggested by Malevich (1969) and Cuzick (1976) and was used in Kedem (1980), Ch. 7, in deriving the asymptotic distribution of $D_{1,N}$. The extension to $D_{k,N}$ is immediate and will not be reproduced here. It follows that when ρ_k is absolutely summable $D_{k,N}$ is asymptotically normal with approximate variance given by (10), and hence approximate 95% probability limits for $D_{k,N}$ are

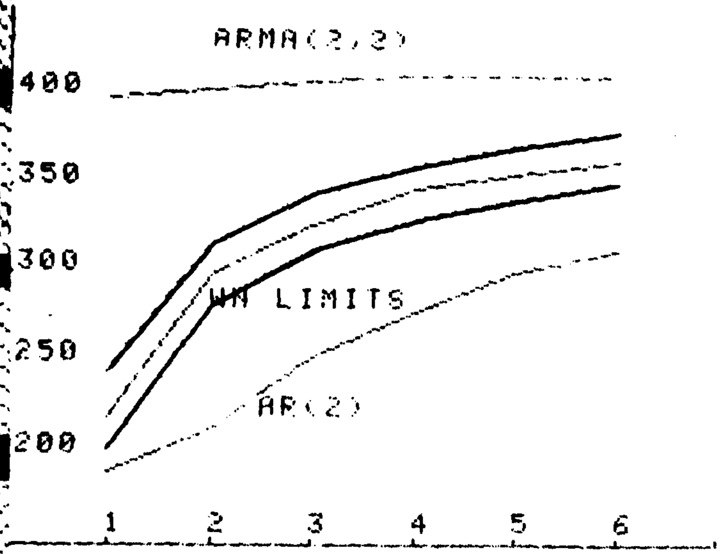
$$(N-1)p^{(k)} \pm 1.96\{\mathrm{Var}\ D_{k,N}\}^{\frac{1}{2}} \tag{12}$$

where $\mathrm{Var}(D_{k,N})$ is given by (10), and $p^{(k)}$ is as in (8). (12) provides a graphical means for assessing the similarity or dissimilarity between processes as expressed by their oscillatory information by way of higher order crossings. Several examples follow.

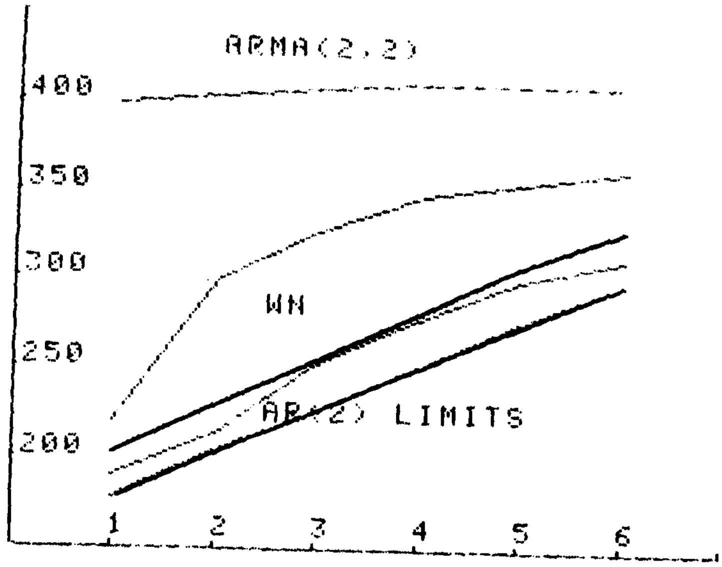
Table 4 gives the probability limits (12) for three processes whose data are given in an appendix in Priestley (1981). It is seen that the HOC obtained from these records with $N = 450$ fall well within the probability limits. The graphical display of the HOC and the limits (12) of these data are given in Figure 1, and show graphically the similarity of the observed data and the hypothesized processes. At the same time the figure shows the dissimilarity or the distance of the particular AR(2) and ARMA(2,2) under consideration from white noise, and also the dissimilarity of these AR(2) and ARMA(2,2). In particular, the rates of increase in the $D_{j,450}$ displayed by the three processes are different and serve as a fast discrimination feature.

Rounded 95% probability limits		Rounded $ED_{j,450}$	$D_{j,450}$
203	246	225	221
282	317	299	299
313	345	329	328
330	361	346	348
342	372	357	355
351	380	365	363
a) White noise: $Z_t = u_t$			
178	204	191	190
205	230	217	214
227	255	241	253
252	282	267	278
277	308	292	300
299	331	315	314
b) $Z_t = 0.4Z_{t-1} - 0.7Z_{t-2} + u_t$			
384	411	397	398
391	415	403	403
394	418	406	407
397	419	408	411
398	421	410	411
400	422	411	411
c) $Z_t = -1.4Z_{t-1} - 0.5Z_{t-2} + u_t - 0.2u_{t-1} - 0.1u_{t-2}$			

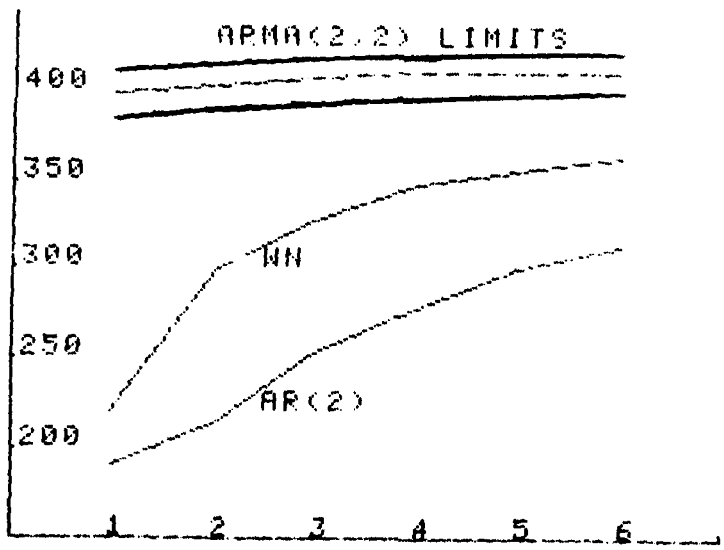
Table 4. Observed $D_{j,450}$ in records given in an Appendix in Priestley (1981). The ARMA(2,2) series c) contains a missing observation which was replaced by 0.0. $u_t \sim N(0,1)$, independent.



a) Observed $D_{j,450}$ from white noise and their theoretical limits (12).



b) Observed $D_{j,450}$ from an AR(2) series and their theoretical limits (12).



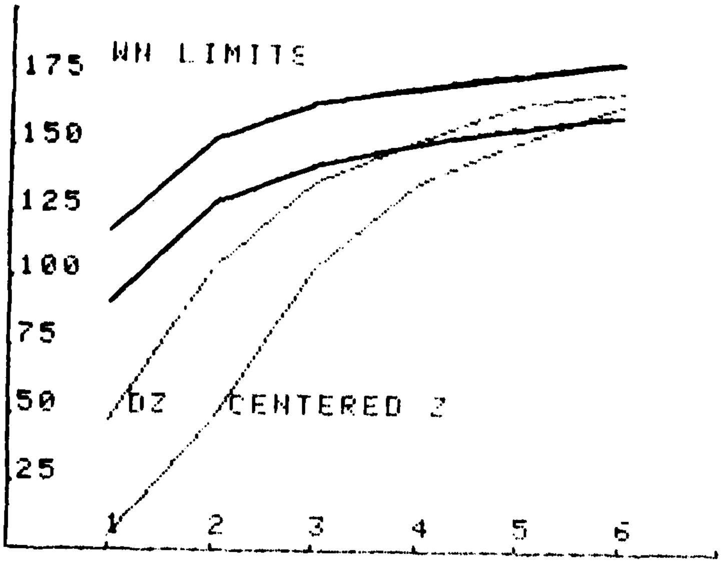
c) Observed $D_{j,450}$ from an ARMA(2,2) series and their theoretical limits (12).

Figure 1. HOC representation for the three series in Table 4. In each case the theoretical limits (12) capture the correct curve obtained from observed $D_{j,450}$ from the indicated series. (The dark lines mark the probability limits.)

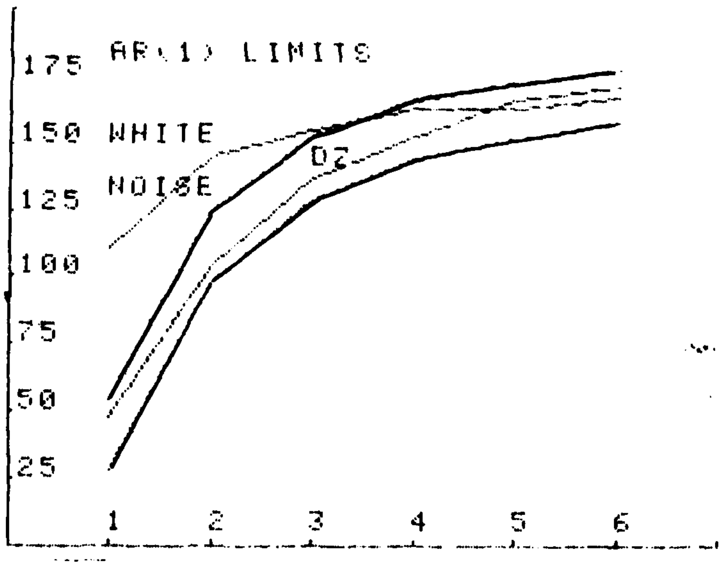
Our next example concerns a HOC diagnostic check for time series models as applied to the temperature series C in Box and Jenkins (1976). Our first step is to see whether the centered series and its first difference are close to being white noise. A quick glance at Figure 2-(a) shows that the series and its difference are far from being white as their HOC do not fall within the white noise limits with $N = 212$. On the other hand, the HOC of the fitted model ($\nabla Z_t = 0.82 \nabla Z_{t-1} + u_t$) suggested by Box and Jenkins (1971), p. 293, are well within the probability limits (12) as seen from Figure 2-(b). The residuals HOC are given in Figure 2-(c) and are shown to fall well within the white noise limits except for $D_{1,212}$ which is still very close to being "in". Thus based on the oscillatory properties of the differenced series and the corresponding residuals it is seen that the fitted model is reasonable.

j	95% limits for white noise		95% limits for AR(1), $\phi = 0.82$		$D_{j,212}$ from $Z_t - \bar{Z}$	$D_{j,212}$ from ∇Z_t	$D_{j,212}$ from the residuals $\nabla Z_t - 0.82 \nabla Z_{t-1}$
1	91	118	28	54	5	48	89
2	129	153	98	125	48	105	139
3	143	166	129	153	105	137	155
4	152	173	144	166	137	153	168
5	158	178	152	173	153	166	172
6	162	182	158	178	166	172	171

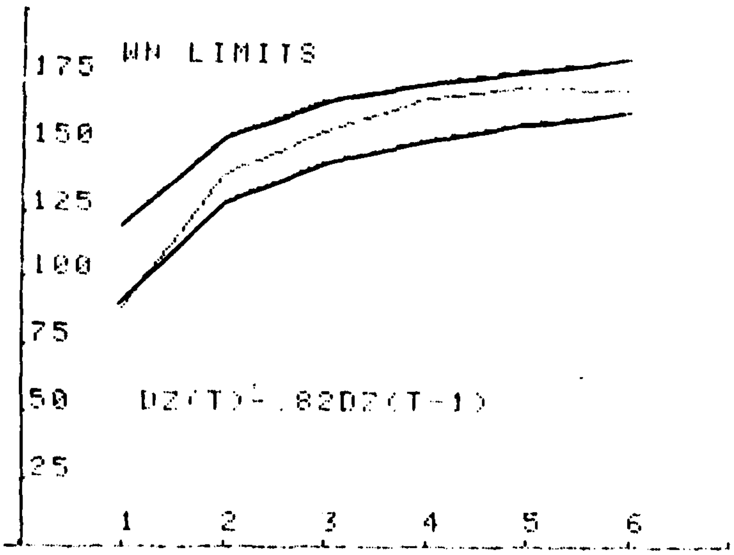
Table 5. Observed HOC from the temperature series C in Box & Jenkins (1976) and probability limits for two models. ($N = 212$). The HOC from ∇Z_t fall within the AR(1) limits. The residuals HOC suggest that the residuals very much resemble white noise.



a) HOC from $Z_t - \overline{Z}$ and ∇Z_t fall outside the white noise limits.



b) HOC from ∇Z_t fall inside the limits (12) under the model $\nabla Z_t = 0.82 \nabla Z_{t-1} + u_t$.



c) HOC of the residuals series resemble very much HOC of white noise.

Figure 2. HOC analysis of the temperature series C in Box & Jenkins (1976) with $N = 212$. (The dark lines mark the probability limits.)

Our last example concerns periodic data plus noise. In many applications a major problem is whether an observed series is made of noise only or of signal plus noise. The oscillation in periodic signals plus noise is usually different from the oscillation displayed by white noise and can be captured very well by HOC. Consider the sunspot series given and analyzed in Anderson (1971). It is well established that this series contains several significant periodic components, and from our point of view it is interesting to measure the deviation of this series from pure white noise as depicted by higher order crossings. By appealing to the limits (12) with $N = 155$ we can see from Figure 3 that the series does indeed contain a "signal" and the hypothesis of white noise is rejected. That is, the sunspot series does not oscillate as white noise.

j	95% probability limits for white noise		$D_{j,155}$ from $Z_t - \bar{Z}$
1	65	89	31
2	93	113	36
3	103	122	81
4	110	128	107
5	114	131	119
6	117	134	121

Table 6. Observed HOC from the sunspot series ($N = 155$) in Anderson (1971) and probability limits for white noise. The hypothesis of noise is rejected.

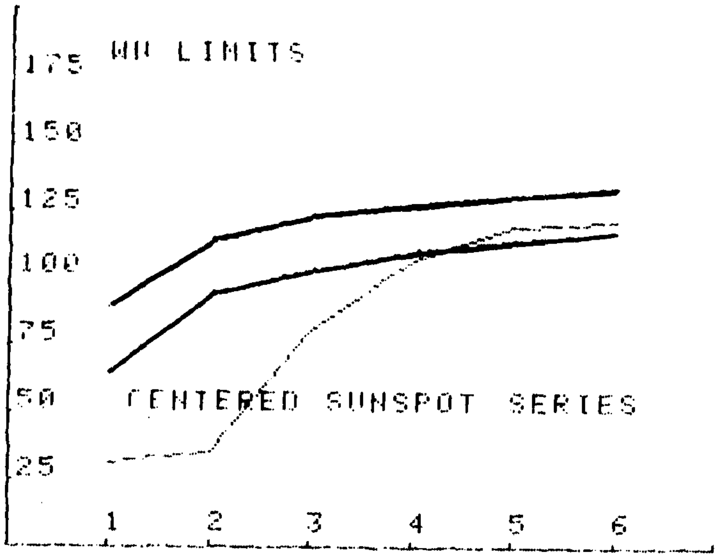


Figure 3. $D_{j,155}$ from $Z_t - \overline{Z}$ derived from the sunspot series. The first four HOC fall outside the white noise limits. (The dark lines mark the probability limits.)

5. Some Power Calculation

In many cases D_1 and D_2 alone are capable of discrimination between time series models. However, there are many examples in which the later D_j do the job while D_1, D_2 fail. In order to illustrate this point we consider several examples of autoregressive moving average models where the hypothesis is that of white noise. At the same time the examples indicate the power of our graphical method obtained from 50 independent series each of length $N = 450$. In each case we bring in the results of 10 realizations in the form of binary row vectors of length 6. Each realization produces $D_{1,450},\ldots,D_{6,450}$ and when a certain $D_{j,450}$ falls outside the probability limits (12) for the white noise case, it is indicated by "1" for success. Otherwise this is indicated by "0". Thus 110011 means that D_1, D_2, D_5, D_6 fall outside the limits (12) for the white noise case, while D_3, D_4 fall inside these limits. Clearly, for series which are not white noise we expect to see many 1's. The closer the process is to white noise the harder is the discrimination problem and the ratio of 0's increases. It is seen from Figure 4 that there are series which are far from being white but their actual axis-crossings behave as those of white noise. On the other hand, the difference in oscillation is captured by higher order crossings as seen in particular from (g) and (i). The power is seen to increase with more significant parameters as is well expected.

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	1	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

a) White noise
power = 0.1

0	1	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	1	0	0	0	0
0	1	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	1	1	0	0	1

b) AR(1), $\phi = 0.05$
power = 0.26

0	0	0	0	0	0
0	1	1	1	0	0
1	0	0	0	0	0
0	0	0	0	0	0
1	0	0	0	0	0
0	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	1	0	0	0	0
0	0	0	0	0	0

c) MA(1), $\theta = 0.1$
power = 0.4

1	0	0	0	0	0
1	1	0	0	0	0
1	0	0	0	0	0
1	1	1	1	1	1
1	1	0	0	0	0
1	1	0	0	0	0
0	1	1	0	0	0
1	1	0	0	0	0
1	0	0	0	0	0
0	0	0	0	0	0

d) AR(1), $\phi = .2$
power = 0.9

1	1	0	0	0	0
1	1	1	0	0	0
1	1	1	0	0	0
1	1	0	0	0	0
1	1	1	1	1	0
1	1	1	1	0	1
1	1	0	1	0	1
1	1	0	0	0	1
1	1	1	1	1	1
1	1	1	0	0	0

e) AR(1), $\phi = 0.5$
power = 1

1	1	1	1	0	1
0	1	0	1	0	0
0	1	1	1	1	1
0	1	0	0	0	0
0	0	0	1	1	1
0	0	1	1	0	1
0	0	0	0	1	1
0	1	1	1	1	1
0	1	1	1	1	0
0	1	1	1	0	0

f) AR(2), $\phi_1 = .1$, $\phi_2 = -0.15$
power = 0.88

0	1	1	1	1	0
0	0	1	1	1	0
0	1	1	1	1	1
0	1	1	1	1	1
0	0	0	0	0	0
0	0	0	0	0	1
0	1	1	0	0	0
0	0	1	1	0	0
0	0	0	1	1	1
0	0	0	1	0	0

g) ARMA(2,2), $\phi_1 = .1$, $\phi_2 = -.2$
$\theta_1 = .2$, $\theta_2 = .1$, power = 0.88

1	1	0	1	0	0
0	1	0	0	0	0
1	0	0	0	0	0
0	0	0	0	0	0
1	0	1	1	1	1
1	1	0	0	1	1
1	0	0	0	0	0
1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1

h) ARMA(1,1), $\phi_1 = .1$, $\theta_1 = -.1$
power = 0.86

0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1

i) ARMA(2,2), $\phi_1 = .1$, $\phi_2 = -.4$, $\theta_1 = 0$
$\theta_2 = .3$, power = 1

Figure 4. Ten typical binary vectors obtained from $D_{j,450}$, $j = 1,\ldots,6$, in relation to the white noise limits from (12). The power was estimated from 50 such binary vectors. 1 means corresponding $D_{j,450}$ is outside the limits. Otherwise 0 is recorded.

References

Anderson, T. W. (1971). The Statistical Analysis of Time Series, New York, Wiley.

Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis Forecasting and control, San Francisco, Holden Day.

Cuzick, J. (1976). A central limit theorem for the number of zeros of a stationary Gaussian process. Annals of Probability, 4, 547-556.

Diananda, P. (1953). Some probability limit theorems with statistical applications. Proc. Cambridge Phil. Soc., 49, 239-246.

Karlin, S. and Taylor, H. M. (1975). A First Course in Stochastic Processes, New York, Academic Press.

Kedem, B. (1980). Binary Time Series, New York, Dekker.

——— (1984). Detection of hidden periodicities by means of higher order crossings, I. Technical Report TR84-55, University of Maryland.

Kedem, B. and Slud, E. (1981). On goodness of fit of time series models: An application of higher order crossings. Biometrika, 68, 551-556.

——— (1982). Time series discrimination by higher order crossings. Annals of Statistics, 10, 786-794.

Malevich, T. L. (1969). Asymptotic normality of the number of crossings of level zero by a Gaussian process. Theory Prob. Applic., 14, 287-295.

Priestley, M. B. (1981). Spectral Analysis and Time Series, New York, Academic Press.

Reed, G. (1983). Some properties and applications of higher order crossings. Ph.D. Thesis, Math. Dept., Univ. of Md.

END

FILMED

10-85

DTIC